R E P O R T R E S U M E S

ED 014 445 SP 001 328

SOME NOTES ON VALIDATING TEACHER SELECTION PROCEDURES.

BY- MEDLEY, DONALD M.
NEW YORK CITY BOARD OF EDUCATION, BROOKLYN, N.Y.

PUB DATE JUN 67

GRANT OEG-1-6-01665-1624
EDRS PRICE MF-$0.25 HC-$0.24 4P.

DESCRIPTORS- *ACHIEVEMENT TESTS, *APTITUDE TESTS, PREDICTION, PREDICTIVE ABILITY (TESTING), TEACHER CHARACTERISTICS, TEACHER EVALUATION, *TEACHER SELECTION, TEST CONSTRUCTION, *TEST SELECTION, TEST VALIDITY,

AT PRESENT, THE PART OF ANY TEACHER EFFECTIVENESS CRITERION THAT CAN BE PREDICTED WITH A SELECTION TEST IS PROBABLY IRRELEVANT TO TEACHER COMPETENCE. TESTING THE VALIDITY OF PREDICTORS OF TEACHER COMPETENCE IS IMPOSSIBLE BECAUSE IT WOULD REQUIRE HIRING A SIZABLE RANDOM SAMPLE OF ALL WHO APPLY FOR POSITIONS, WITHOUT PRIOR SCREENING. FURTHER, TEACHER APTITUDE TESTS WRONGLY ASSUME THAT THE FACTORS IN SUCCESSFUL TEACHING OPERATE PRIOR TO THE START OF TEACHING IN A SPECIFIC SCHOOL OR SCHOOL SYSTEM. INSTEAD, TEACHING BEHAVIOR PROBABLY VARIES WITH THE TEACHING SITUATION. ACCORDINGLY, AN ACHIEVEMENT RATHER THAN A PREDICTIVE MODEL SHOULD BE USED IN HIRING NEW TEACHERS. FAST LEARNING (E.G., AS MEASURED BY COLLEGE GRADES) IS ONE SUCH MEASURE. TEACHING PERFORMANCE IS ANOTHER ACHIEVEMENT MEASURE. THAT IS, IF, AFTER ONE YEAR, THE PROBATIONARY TEACHER HAS NOT LEARNED TO TEACH, HE SHOULD NOT BE REHIRED. THIS DOCUMENT APPEARED IN GILBERT, H.B., AND LANG, G., "TEACHER SELECTION METHODS" NEW YORK, 1967. (RD)

LB 2328
.C61

TEACHER SELECTION METHODS

Project No. 6-1665
Grant No. OEG 1-6-061665-1624

Harry B. Gilbert

Pennsylvania State University

Gerhard Lang

Montclair State College

June 1967

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED014445

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

BOARD OF EXAMINERS
Board of Education of
The City of New York

New York, New York

Some Notes on Validating Teacher Selection Procedures

Donald M. Medley

Educational Testing Service

As far as I know, all teacher selection procedures presently used are based on the model of aptitude testing; in other words, in constructing selection instruments the effort has been to devise a battery which would predict success on the job. Validation studies have, accordingly, sought to establish predictive validity against some kind of a criterion measure of teacher effectiveness obtained after the teacher has been admitted to employment.

A selection battery that could do this job fairly well would be useful indeed; it would enable the selection agency to compile a list of candidates with the candidate who would make the best teacher at the top and the one who would make the poorest at the bottom. Then it would be possible to appoint as many teachers as were needed in a given year, beginning at the top of the list, knowing that the best possible set of candidates had been chosen. This is a beautiful ideal; but it just will not work in practice. To my mind, the sooner all attempts to validate selection procedures in this way are abandoned the better.

In the first place, when you consider the nature of what you are trying to predict--teacher competence--it seems highly improbable that it can ever be measured with a paper-and-pencil test, or any other device which could conceivably be used on the scale necessary for teacher selection in large cities. There is considerable experimental data which confirms this pessimistic point of view. Most of the predictive validities obtained in studies done in the past have been below .30; very few have exceeded .40. And the improvement in predictive efficiency obtained with such small correlations is practically negligible.

In the second place, even if the correlations obtained were large enough to improve selection, their value would be suspect because of the limited validity of the criteria on which they would have to be based. At the present state of the art of measuring teacher competence, it is fair to say that the part of any teacher effectiveness criterion we can predict with a selection test is probably irrelevant to teacher competence anyhow.

Finally, it is extremely difficult actually to carry out a validity study based on this model, since to do so requires that--

for experimental purposes--a good sized random sample of all who apply for positions in a school system be admitted to teaching without any kind of selection. This is probably illegal, but in any case is not practical, in most large school systems.

A more fundamental reason for abandoning the aptitude test model is the fact that its use is based on an untenable assumption; the assumption that the major factors which determine whether a candidate will succeed or fail as a teacher operate before he enters the school system. Such things as what kind of a school and neighborhood the teacher is assigned to, the characteristics of his pupils and the facilities and materials available to him when he tries to teach them, and the amount and kind of support he receives from his peers and superiors in the school system are seen as distinctly less important than such things as what college courses he has had and what he learned in them, whether he has worked in summer camps during his undergraduate days, and how happy his childhood was, in determining his future as a teacher. This assumption is clearly implied by a conception of the selection problem as one of identifying among applicants for positions those predestined to become competent teachers.

It seems more realistic not to assume that the future of any of the candidates has been (or should have been) decided at the time when the selection takes place, but only that the candidates will vary in the degree to which they have mastered that part of their preparation which can be obtained before they enter the system. The selection problem would then be seen as one of assessing past learning rather than one of predicting future performance.

The problem of validating a selection battery would then become a matter of content validity rather than of predictive validity; and this kind of validity is much more likely to be achieved by a paper-and-pencil test or one of the other techniques which a practical selection battery is likely to contain. From this point of view it may be said that teacher selection should be based on an achievement model rather than a prediction model.

After a teacher has been admitted to probationary status in the school system, we are still faced with a problem of eliminating those candidates who have satisfactorily completed their preservice training but cannot teach successfully. This could be regarded as a problem in prediction; but I prefer to say that it is another problem in achievement testing. The first years of a teacher's career should be viewed as a part of his training; if by the end of his probationary period he has not learned to teach, he is not ready to be admitted to permanent tenure and should probably be let go.

If proven ability to teach were made a prerequisite for a permanent appointment as a teacher, some kind of a teaching test administered during the probationary period would become a part of the final selection procedure. This might be a formal affair involving the teaching of a standard lesson under observation, or an informal one involving ratings by supervisors. In either case, it should be a much fairer and more valid selection instrument than anything that could be used at the pre-service level.

If we adopt the achievement model for validating selection instruments as I have suggested, how can we find out whether our overall selection policies are achieving their ultimate purpose of improving the quality of teaching in the schools?

There is one approach to this problem which (so far as I know) has never been tried out. This approach would involve the continuous and routine monitoring of the quality of teachers in the school system. An analogy to quality control on an industrial production line may be drawn here. In industry, a sample of the objects being manufactured is removed from the production line at certain intervals for inspection. Such an inspection is much too thorough and expensive to be applied to every item which comes off the line; but if objects so examined are randomly selected, it is possible to obtain a quite satisfactory estimate of the average quality of the entire output by inspecting only a sample of it.

In similar fashion, the city school system could select a stratified random sample of all the teachers who earn tenure in the system each year, and make a thorough and careful study of each teacher <u>after he has received tenure</u>. The results of the study as related to any one individual would be kept confidential, and guarantees made that they would not affect his career in the system; but group data would be used to make inferences about the teacher population as a whole. Such data would provide precise information about the overall quality of teachers brought into the schools year by year, so that important trends could be detected; they would also provide diagnostic information about areas of strength and weakness on the basis of which changes could be made.

The fact that only group scores would be used would make it possible to employ instruments and procedures whose reliabilities are too low to be useful for individual diagnosis; and the fact that the number of teachers tested would be small relative to the total number employed in the system would make it possible to use relatively expensive and time-consuming procedures. Although we do not presently have the capability of measuring the competence of individual teachers reliably and economically, I am confident that we could reliably estimate the <u>average</u> competence of all the teachers

in a system, using techniques already available to us, at a per teacher cost which would be quite reasonable.

I have introduced this idea of quality control of teachers in a system in the context of the problem of monitoring teacher selection policies and procedures, because that is our immediate concern here; but I would point out the obvious fact that such information would have many other uses, some of which might be regarded as more important than this one. I see no reason why all of these purposes should not be achieved at the same time. Nor do I see any other feasible way of achieving the one we are concerned with--that of assessing the effects of selection policies and practices as they are used.